

Webbasierte Suche in Vorlesungsvideos auf Basis der Transkripte eines Spracherkenners

Stephan Repp, Andreas Groß, Christoph Meinel

Hasso Plattner Institut (HPI), Universität Potsdam
Prof.-Dr.-Helmert-Str. 2-3
D-14482 Potsdam
{repp, gross, meinel}@hpi.uni-potsdam.de

Abstract: Immer mehr Hochschulen nutzen die Möglichkeit, ihre Vorlesungen als Videos aufzuzeichnen, in multimedialen Datenbanken abzulegen und für den Zugriff über das Internet bereitzustellen. Bislang ist die detaillierte Suche in diesen Multimedia-Daten jedoch nur begrenzt möglich. Neben der Auswertung der Präsentationsfolien ist ein weiterer vielversprechender Ansatz für die Suche die Verwendung der Sprachtranskripte des Sprechers. Die Ergebnisse eines Spracherkenners (SRE) werden dazu genutzt, um in den Vortragsaufzeichnungen nach Schlüsselwörtern und deren semantischen Zusammenhängen zu suchen. Das Benutzerschnittstelle erlaubt darüber hinaus ein einfaches dynamisches Navigieren innerhalb der einzelnen Vorlesung.

Vorlesungen werden in Form von Multimediadaten aufgezeichnet. Mit Hilfe eines Spracherkenners wird die Sprache des Sprechers in eine Transkription überführt. Nach einer standardisierten Vorverarbeitung werden die einzelnen Terme mit ihrer entsprechenden Zeitmarke in eine Datenbank abgelegt. Ausgehend von diesen Daten wird nun ein Index von zusammenhängenden Bereichen erstellt [RM06b, RM06a]. Ein Bereich (Chain = Verkettung) ist ein Stelle im Video, bei der ein Term oft vom Sprecher erwähnt wird. Ein Chain beginnt mit der ersten Erwähnen und endet beim letzten Auftreten des Terms. Vergeht zwischen dem Auftreten der Terme eine längere Zeit, so wird ein neuer Chain

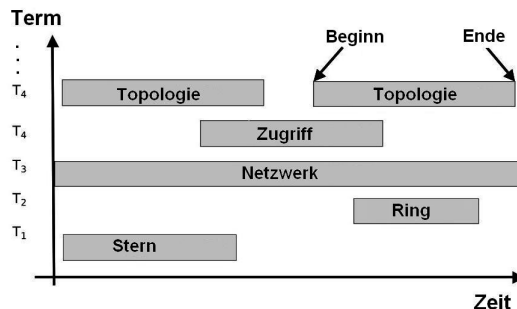


Abbildung 1: Chains

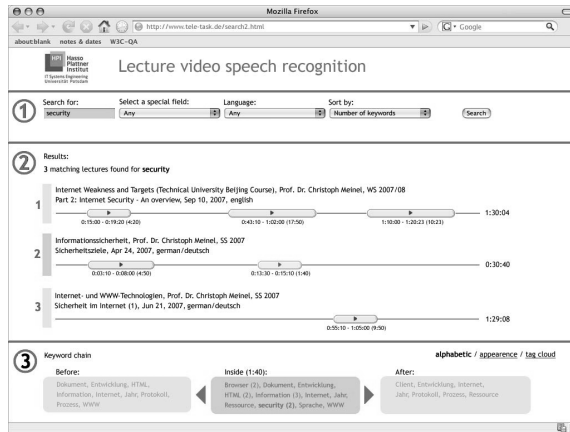


Abbildung 2: Benutzerschnittstelle

angelegt (Abbildung 1). Aus diesem häufigen Auftreten des Terms in der Vorlesung kann nun geschlossen werden, dass der Term an dieser Stelle der Vorlesung relevant ist.

Mit dieser Indexierung können nun Mehrdeutigkeiten aufgelöst werden. In Abbildung 1 wird der Term *Topologie* im Zusammenhang mit einem *Netzwerk* erwähnt und der Term *Ring* wird ebenso im Kontext zu *Topologie* vom Sprecher verwendet. Der Sinn des Wortes *Ring* kann nun aus dem Zusammenhang erschlossen werden. Darüber hinaus, liefert der Chain-Index einen Index für jedes Videosegment. Auf den Suchstring *Netzwerk* erhält der Benutzer nun die relevante Stelle sowie die innerhalb liegenden Chains (*Ring*, *Stern*, *Topologie*). Die innerhalb liegenden Chains repräsentieren nun den Inhalt des Video-Segments. Ebenso hat jedes Chain einen Bereich *davor*, einen *nachfolgenden* und einen *übergeordneten* Bereich (Abbildung 2). Der Benutzer hat nun die Möglichkeit in diesen Bereichen zu navigieren um den semantisch passenden Video-Bereich zu finden.

Das vorgestellte System erlaubt das Browsen in einem Archiv von Vorlesungsmitschnitten. Das Ergebnis zeigt, dass es möglich ist, Schlüsselwörter zum Ergebnisset hinzuzufügen, so dass Studenten die gesuchten Informationen leichter finden können.

Literatur

- [RM06a] Stephan Repp und Christoph Meinel. Segmenting of Recorded Lecture Videos - The Algorithm VoiceSeg. In *Proceedings of the 11th Signal Processing and Multimedia Applications*, Seiten 317–322, August 2006.
- [RM06b] Stephan Repp und Christoph Meinel. Semantic Indexing for Recorded Educational Lecture Videos. In *4th IEEE Conference on Pervasive Computing and Communications Workshops (PerCom 2006 Workshops)*, 13-17 March 2006, Pisa, Italy, Seiten 240–245. IEEE Computer Society, 2006.